**2007/17**

■

# Lattice based extended formulations for integer linear equality systems

Karen Aardal and Laurence A. Wolsey

**CORE**
Voie du Roman Pays 34
B-1348 Louvain-la-Neuve, Belgium.
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
E-mail: corestat-library@uclouvain.be
http://www.uclouvain.be/en-44508.html

CORE DISCUSSION PAPER
2007/17

# Lattice based extended formulations for integer linear equality systems

Karen AARDAL[1] and Laurence A. WOLSEY[2]

March 2007

## Abstract

We study different extended formulations for the set $X = \{\mathbf{x} \in \mathbf{Z}^n \mid \mathbf{Ax} = \mathbf{Ax}^0\}$ in order to tackle the feasibility problem for the set $X_+ = X \cap \mathbf{Z}^n_+$. Here the goal is not to find an improved polyhedral relaxation of conv($X_+$), but rather to reformulate in such a way that the new variables introduced provide good branching directions, and in certain circumstances permit one to deduce rapidly that the instance is infeasible. For the case that $\boldsymbol{A}$ has one row $\boldsymbol{a}$ we analyze the reformulations in more detail. In particular, we determine the integer width of the extended formulations in the direction of the last coordinate, and derive a lower bound on the Frobenius number of $\boldsymbol{a}$. We also suggest how a decomposition of the vector $\boldsymbol{a}$ can be obtained that will provide a useful extended formulation. Our theoretical results are accompanied by a small computational study.

**Keywords**: integer programming feasibility, integer width, branching directions, reduced lattice bases, Frobenius number.

---

[1] Centrum voor Wiskunde en Informatica, Amsterdam and the Eindhoven Institute of Technology, The Netherlands. E-mail: aardal@cwi.nl

[2] CORE and INMA, Université catholique de Louvain, Belgium. E-mail: Laurence.wolsey@uclouvain.be

The work was partly carried out within the framework of ADONET, a European network in Algorithmic Discrete Optimization, contract no MRTN-CT-2003-504438. The first author is financed in part by the Dutch BSIK/BRICKS project. The research was carried out in part while the second author visited CWI, Amsterdam with the support of the NWO visitor grant number B 61-556.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

# 1 Introduction

Recently, several hard integer programming feasibility problems have been successfully tackled using a lattice reformulation proposed by Aardal et al. [2, 3]. These problems include various equality knapsacks of Cornuéjols et al. [7], Aardal and Lenstra [4], market split instances of Cornuéjols and Dawande [6, 1], and financial planning instances of Louveaux and Wolsey [14].

The approach proposed earlier is to replace the feasibility problem for the set $X = \{\boldsymbol{x} \in \mathbb{Z}^n_+ \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$ by a feasibility problem over an extended formulation for the set, namely the set $\{(\boldsymbol{x}, \boldsymbol{\mu}) \in \mathbb{Z}^n_+ \times \mathbb{Z}^{n-m} \mid \boldsymbol{x} = \boldsymbol{x}^0 + \boldsymbol{Q}\boldsymbol{\mu}\}$ where $\boldsymbol{Q}$ is a reduced basis of the lattice $\{\boldsymbol{y} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{y} = \boldsymbol{0}\}$. It turns out that both a special purpose branch-and-bound code and commercial mixed integer programming solvers are then much more successful at solving the test instances with the extended formulation than with the original formulation. In related work Aardal and Lenstra [4] have considered knapsack sets $\{\boldsymbol{x} \in \mathbb{Z}^n_+ \mid \boldsymbol{a}\boldsymbol{x} = b\}$ where $\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2$, $M_1, M_2 \in \mathbb{Z}_{>0}$, $\boldsymbol{p}^1, \boldsymbol{p}^2 \in \mathbb{Z}^n$, where the vectors $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ are short compared to the length of $\boldsymbol{a}$. For such instances they have used the reduced basis approach to detect a good direction for branching and to demonstrate infeasibility for large values of $b$, and they theoretically derive a strong lower bound on the Frobenius number of $\boldsymbol{a}$ in the special case when $M_2 = 1$ and $\boldsymbol{p}^1 \in \mathbb{Z}^n_{>0}$.

Here we pursue this lattice viewpoint. In Section 2 we show that the formulations presented above are just two extremes of a family of extended formulations of the original set $X$ taking the form $\{(\boldsymbol{x}, \boldsymbol{\mu}) \in \mathbb{Z}^n \times \mathbb{Z}^s \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}^0 + \boldsymbol{T}\boldsymbol{\mu}\}$ with $s$ additional variables and $m + s$ constraints with $0 \leq s \leq n - m$, where each row of $\boldsymbol{A}$ can be expressed as an integer multiple of the rows of $\boldsymbol{P}$. Note that whereas in polyhedral combinatorics extended formulations are typically used to provide better polyhedral approximations of $conv(X)$, here the extended formulations do not change the underlying polyhedron. However by isolating good branching directions, the formulations are made more effective for treatment by branch-and-bound.

In Section 3, we consider the special case of knapsack sets when $m = 1$ and $\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2$. For the reformulation with two equations and one additional variable ($s = 1$), we calculate the integer width of the underlying polyhedron in the direction of the auxiliary variable $\mu$. This in turn leads us to a lower bound on the Frobenius number, simplifying and generalizing the earlier proof from [4] cited above.

In Section 4.1 we make precise how hidden structure in $\boldsymbol{A}$ can be detected by reduced basis calculations, and also how this leads to interesting reformulations. In Section 4.2 we demonstrate computationally examples of hidden structure, we test how effective the reformulations are in solving feasibility problems, and we examine the quality of the lower bounds on the Frobenius number.

## 1.1 Preliminaries

In the following we will frequently refer to lattices and reduced lattice bases. We will define and describe these briefly here. For a more detailed exposition, see for instance Cassels [5], Kannan [8], Lenstra [11, 12], and Lovász [13].

Let $\boldsymbol{b}^1, \ldots, \boldsymbol{b}^l$ be linearly independent vectors in $\mathbb{R}^n$. The set

$$L = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \sum_{j=1}^{l} \lambda_j \boldsymbol{b}^j,\ \lambda_j \in \mathbb{Z},\ 1 \leq j \leq l\}$$

is called a *lattice*. The set of vectors $\{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^l\}$ is called a *lattice basis*. The lattice generated by the basis $\boldsymbol{B}$ is denoted by $L(\boldsymbol{B})$. If $\boldsymbol{B}$ is clear from the context we write just $L$.

Suppose $\boldsymbol{A}$ is an integer $m \times n$ matrix of full row rank. Then $N(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, i.e., $N(\boldsymbol{A})$ is the null-space of $\boldsymbol{A}$. We use the notation $L(N(\boldsymbol{A}))$ to denote the lattice that is generated by a basis of all *integer vectors* in $N(\boldsymbol{A})$.

The *rank* of a given lattice $L$, rk $L$, is equal to the dimension of the Euclidean vector space generated by a basis of $L$.

Let $\boldsymbol{B} = (\boldsymbol{b}^1, \ldots, \boldsymbol{b}^l)$ be a basis for a given lattice $L$, and let $\boldsymbol{b}^{1*}, \ldots, \boldsymbol{b}^{n*}$ be the associated Gram-Schmidt vectors, which are derived recursively as follows.

$$\boldsymbol{b}^{1*} = \boldsymbol{b}^1 \text{ and } \boldsymbol{b}^{j*} = \boldsymbol{b}^j - \sum_{k=1}^{j-1} \mu_{jk} \boldsymbol{b}^{k*}, \quad 2 \leq j \leq l,$$

where

$$\mu_{jk} = ((\boldsymbol{b}^j)^T \boldsymbol{b}^{k*})/(||\boldsymbol{b}^{k*}||^2), \quad 1 \leq k < j \leq l\,.$$

The basis $\boldsymbol{B} = (\boldsymbol{b}^1, \ldots, \boldsymbol{b}^l)$ of the lattice $L$ is *reduced in the sense of Lenstra-Lenstra-Lovász* [10] if the following holds:

$$|\mu_{jk}| \leq \frac{1}{2} \quad \text{for } 1 \leq k < j \leq l,$$

$$||\boldsymbol{b}^{j*} + \mu_{j,j-1} \boldsymbol{b}^{(j-1)*}||^2 \geq \frac{3}{4} ||\boldsymbol{b}^{(j-1)*}||^2 \quad \text{for } 1 < j \leq l.$$

Such a basis can be found in polynomial time using the so-called $LLL$ algorithm [10].

We will also need to find a basis $\boldsymbol{Q}$ of $L(N(\boldsymbol{A}))$ and a vector $\boldsymbol{x}^0$ satisfying $\boldsymbol{A}\boldsymbol{x}^0 = \boldsymbol{b}$, if such a vector exists. They can be found by reducing the basis of the lattice defined by the columns of the $(n + m + 1) \times (n + 1)$-matrix

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & N_1 \\ N_2\boldsymbol{A} & -N_2\boldsymbol{b} \end{pmatrix},$$

where $N_1$ and $N_2$ are appropriately chosen positive integers, see [3]. The first $n-m+1$ columns of the reduced basis $\boldsymbol{B}'$ will be

$$\begin{pmatrix} \boldsymbol{Q} & \boldsymbol{x}^0 \\ \boldsymbol{0} & \pm N_1 \\ \boldsymbol{0} & \boldsymbol{b} \end{pmatrix}. \tag{1}$$

If there does not exists vector $\boldsymbol{x}^0$ satisfying $\boldsymbol{A}\boldsymbol{x}^0 = \boldsymbol{b}$, then element $b'_{n+1,n-m+1}$ of $\boldsymbol{B}'$ will be equal to $\pm kN_1$ for some integer $k > 1$, indicating that there exists a vector $\boldsymbol{x}$ satisfying $\boldsymbol{A}\boldsymbol{x} = k\boldsymbol{b}$.

The *Hermite Normal Form* of a matrix $\boldsymbol{A} \in \mathbb{Z}^{m\times n}$ of full row rank, HNF($\boldsymbol{A}$), is obtained by multiplying $\boldsymbol{A}$ by an $n \times n$ unimodular matrix $\boldsymbol{U}$ to obtain the form $(\boldsymbol{D},\ \boldsymbol{0})$, where $\boldsymbol{D} \in \mathbb{Z}^{m\times m}$ is a nonsingular, nonnegative lower triangular matrix with the unique row maximum along the diagonal. The Hermite Normal Form can be calculated in polynomial time using for instance the *LLL* algorithm, see [15], or that of Kannan and Bachem [9]. Note that when $\boldsymbol{D} \neq \boldsymbol{I}$, the matrix $\boldsymbol{D}^{-1}\boldsymbol{A}$ is integral and HNF($\boldsymbol{D}^{-1}\boldsymbol{A}$) = $(\boldsymbol{I},\ \boldsymbol{0})$.

We will need to use one property concerning projections of equality sets over the reals.

**Lemma 1** *Let $\boldsymbol{C}$ and $\boldsymbol{D}$ be rational matrices of dimension $m \times n$ and $m \times p$ respectively, and let $\boldsymbol{b}$ be a rational $m$-vector. If*

$$T = \{\boldsymbol{x} \in \mathbb{R}^n \mid \text{ there exists } \boldsymbol{y} \in \mathbb{R}^p \text{ with } \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y} = \boldsymbol{b}\},$$

*and if $\Delta$ is a basis of $N(\boldsymbol{D}^T)$, then*

$$T = \{\boldsymbol{x} \in \mathbb{R}^n \mid \Delta^T\boldsymbol{C}\boldsymbol{x} = \Delta^T\boldsymbol{b}\}.$$

**Proof:** Multiplying by $\Delta^T$, it is immediate that $\{\boldsymbol{x} \in \mathbb{R}^n \mid$ there exists $\boldsymbol{y} \in \mathbb{R}^p$ with $\boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y} = \boldsymbol{b}\} \subseteq \{\boldsymbol{x} \in \mathbb{R}^n \mid \Delta^T\boldsymbol{C}\boldsymbol{x} = \Delta^T\boldsymbol{b}\}$ as $\Delta^T\boldsymbol{D} = 0$. If the inclusion is strict, there exists $\boldsymbol{x}^* \in \mathbb{R}^n$ such that $\Delta^T\boldsymbol{C}\boldsymbol{x}^* = \Delta^T\boldsymbol{b}$, but $\{\boldsymbol{y} \in \mathbb{R}^p \mid \boldsymbol{D}\boldsymbol{y} = \boldsymbol{b} - \boldsymbol{C}\boldsymbol{x}^*\} = \emptyset$. This implies the existence of a vector $\boldsymbol{u}$ such that $\boldsymbol{u}\boldsymbol{D} = 0$ and $\boldsymbol{u}(\boldsymbol{b} - \boldsymbol{C}\boldsymbol{x}^*) \neq 0$. Howeever as $\Delta$ is a basis of $N(\boldsymbol{D}^T)$, there exists $\boldsymbol{v}$ such that $\boldsymbol{u} = \boldsymbol{v}\Delta^T$. But now $\boldsymbol{u}(\boldsymbol{b} - \boldsymbol{C}\boldsymbol{x}^*) = \boldsymbol{v}\Delta^T(\boldsymbol{b} - \boldsymbol{C}\boldsymbol{x}^*) = 0$, a contradiction. □

The set of nonnegative (positive) integers in $\mathbb{R}^n$ is denoted by $\mathbb{Z}^n_+$ ($\mathbb{Z}^n_{>0}$). Similar notation is used for the nonnegative real numbers.

## 2 Formulations for equality constrained integer programs

Our goal in this section is to derive reformulations of the integer programming equality set

$$X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$$

and the associated affine set

$$Y = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\},$$

where $\boldsymbol{A} \in \mathbb{Z}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{Z}^n$. Later these will be used in testing feasibility or optimizing over the set $X^+ = X \cap \mathbb{Z}^n_+$.

Throughout the paper we will assume that $X \neq \emptyset$. This can be tested by calculating HNF($\boldsymbol{A}$). Note that for any point $\boldsymbol{x}^0 \in X$, we can then write $X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$.

We will be dealing with the kernel lattice $L(N(\boldsymbol{A})) = \{\boldsymbol{y} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{y} = \boldsymbol{0}\}$. We assume that rank$(\boldsymbol{A}) = m$, so rk $(L(N(\boldsymbol{A}))) = n - m$.

The following is immediate as $\boldsymbol{x} \in X$ if and only if $\boldsymbol{x} - \boldsymbol{x}^0 \in L(N(\boldsymbol{A}))$.

**Observation 2** *Let $\boldsymbol{Q} \in \mathbb{Z}^{n\times(n-m)}$ be a lattice basis of $L(N(\boldsymbol{A}))$. Then*

$$X = \{\boldsymbol{x}^0\} + L(N(\boldsymbol{A})) = \{\boldsymbol{x} \mid \boldsymbol{x} = \boldsymbol{x}^0 + \boldsymbol{Q}\boldsymbol{\mu},\ \boldsymbol{\mu} \in \mathbb{Z}^{n-m}\}.$$

Now we derive a larger family of reformulations of the set $X$. Specifically suppose that two matrices $\boldsymbol{P} \in \mathbb{Z}^{(m+s)\times n}$ and $\boldsymbol{M} \in \mathbb{Z}^{m\times(m+s)}$ are given satisfying $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{P}$. Two natural questions are:

1. Under what conditions does there exist an extended formulation for $X$ of the form
$$X = \{x \in \mathbb{Z}^n \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}^0 + \boldsymbol{T}\boldsymbol{\mu}, \boldsymbol{\mu} \in \mathbb{Z}^s\}$$
where $\boldsymbol{T} \in \mathbb{Z}^{(m+s)\times s}$?

2. How should one choose the matrices $\boldsymbol{P}, \boldsymbol{M}, \boldsymbol{T}$ so as to obtain "interesting" extended formulations?

We will use the following notation:

$$U^{\boldsymbol{P},\boldsymbol{T}} = \{(\boldsymbol{x}, \boldsymbol{\mu}) \in \mathbb{Z}^n \times \mathbb{Z}^s \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}^0 + \boldsymbol{T}\boldsymbol{\mu}\} \text{ and } X^{\boldsymbol{P},\boldsymbol{T}} = \text{proj}_{\boldsymbol{x}} U^{\boldsymbol{P},\boldsymbol{T}}.$$

$$V^{\boldsymbol{P},\boldsymbol{T}} = \{(\boldsymbol{x}, \boldsymbol{\mu}) \in \mathbb{R}^n \times \mathbb{R}^s \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}^0 + \boldsymbol{T}\boldsymbol{\mu}\}, \text{ and } Y^{\boldsymbol{P},\boldsymbol{T}} = \text{proj}_{\boldsymbol{x}} V^{\boldsymbol{P},\boldsymbol{T}}.$$

Here, $\text{proj}_{\boldsymbol{x}} U^{\boldsymbol{P},\boldsymbol{T}} = \{\boldsymbol{x} \in \mathbb{Z}^n \mid$ there exists $\boldsymbol{\mu} \in \mathbb{Z}^s$ with $(\boldsymbol{x}, \boldsymbol{\mu}) \in U^{\boldsymbol{P},\boldsymbol{T}}\}$. In this notation $X = X^{\boldsymbol{A},\boldsymbol{0}}$ and Observation 2 states that $X = X^{\boldsymbol{I},\boldsymbol{Q}}$ when $\boldsymbol{Q}$ is a lattice basis of $N(L(\boldsymbol{A}))$.

In Propositions 3 and 4 below we address the first question, and derive conditions ensuring that $X = X^{\boldsymbol{P},\boldsymbol{T}}$.

**Proposition 3** *Suppose that $\boldsymbol{A}$, $\boldsymbol{P}$ and $\boldsymbol{M}$ are given with $\boldsymbol{P} \in \mathbb{Z}^{(m+s)\times n}$, $\boldsymbol{M} \in \mathbb{Z}^{m\times(m+s)}$ and $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{P}$. Then*

$$X = X^{\boldsymbol{P},\boldsymbol{T}}$$

*if $\boldsymbol{T} \in \mathbb{Z}^{(m+s)\times s}$ is a lattice basis of $L(N(\boldsymbol{M}))$.*

**Proof:** If $\boldsymbol{x} \in X^{\boldsymbol{P},\boldsymbol{T}}$, then there exists $\boldsymbol{\mu}$ such that $(\boldsymbol{x}, \boldsymbol{\mu}) \in U^{\boldsymbol{P},\boldsymbol{T}}$. Left multiplying by $\boldsymbol{M}$, we see that $\boldsymbol{MPx} = \boldsymbol{MPx}^0 + \boldsymbol{MT\mu} = \boldsymbol{MPx}^0$, or $\boldsymbol{Ax} = \boldsymbol{Ax}^0$. So $X^{\boldsymbol{P},\boldsymbol{T}} \subseteq X$.

Conversely $\boldsymbol{x} \in X$ implies that $\boldsymbol{Ax} = \boldsymbol{Ax}^0$, or that $\boldsymbol{MP}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{0}$. As $\boldsymbol{P}(\boldsymbol{x} - \boldsymbol{x}^0) \in \mathbb{Z}^{m+s}$ and $\boldsymbol{T}$ is a lattice basis of $L(N(\boldsymbol{M}))$, $\boldsymbol{P}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{T\mu}$ for some $\boldsymbol{\mu} \in \mathbb{Z}^s$. Thus $X \subseteq X^{\boldsymbol{P},\boldsymbol{T}}$. □

Note that a slightly stronger statement can be made.

**Proposition 4** *If the Hermite Normal Form of $\boldsymbol{P}$ is of the form $(\boldsymbol{I}, \boldsymbol{0})$, then*

$$X = X^{\boldsymbol{P},\boldsymbol{T}}$$

*if and only if $\boldsymbol{T} \in \mathbb{Z}^{(m+s)\times s}$ is a lattice basis of $L(N(\boldsymbol{M}))$.*

**Proof:** Necessarily if $X = X^{\boldsymbol{P},\boldsymbol{T}}$, then the columns of $\boldsymbol{T}$ lie in $L(N(\boldsymbol{M}))$. If they do not form a lattice basis, there exists $\boldsymbol{t}^* \in L(N(\boldsymbol{M}))$ that is not in the sublattice $L(\boldsymbol{T})$. Also, as $\mathrm{HNF}(\boldsymbol{P}) = (\boldsymbol{I}, \boldsymbol{0})$, there exists $\boldsymbol{x}^*$ such that $\boldsymbol{Px}^* = \boldsymbol{t}^*$. Now the point $\boldsymbol{y}^* = \boldsymbol{x}^0 + \boldsymbol{x}^*$ lies in $X$ as $\boldsymbol{Ay}^* - \boldsymbol{Ax}^0 = \boldsymbol{Ax}^* = \boldsymbol{MPx}^* = \boldsymbol{Mt}^* = \boldsymbol{0}$, but $\boldsymbol{P}(\boldsymbol{y}^* - \boldsymbol{x}^0) = \boldsymbol{Px}^* = \boldsymbol{t}^* \neq \boldsymbol{T\mu}$ for any $\boldsymbol{\mu} \in \mathbb{Z}^s$, and thus $X \neq X^{\boldsymbol{P},\boldsymbol{T}}$. □

Here we address the second question specifying one way to choose appropriate matrices $\boldsymbol{P}$ and $\boldsymbol{T}$. The more restrictive conditions imposed will be used in Sections 3 and 4.1 to make "good" choices using reduced bases.

Consider a sublattice $L'$ of $L(N(\boldsymbol{A}))$ of rank $0 \leq r \leq n - m$ with lattice basis $\boldsymbol{R} \in \mathbb{Z}^{n\times r}$, and $\boldsymbol{Q} = (\boldsymbol{R}, \boldsymbol{S})$ an extension to a basis of $L(N(\boldsymbol{A}))$. Let $\boldsymbol{P}^T \in \mathbb{Z}^{n\times n-r}$ be a lattice basis of $L(N(\boldsymbol{R}^T)) = \{\boldsymbol{y} \in \mathbb{Z}^n \mid \boldsymbol{R}^T\boldsymbol{y} = \boldsymbol{0}\}$, and $s = n - m - r$.

With $\boldsymbol{P}$ constructed in this way, we have that:

i) $\boldsymbol{A} = \boldsymbol{MP}$ for some $m \times (m+s)$ integer matrix $\boldsymbol{M}$. $\boldsymbol{R} \subseteq \boldsymbol{Q}$ implies $L(N(\boldsymbol{Q}^T)) \subseteq L(N(\boldsymbol{R}^T))$, and as $\boldsymbol{A}^T \in L(N(\boldsymbol{Q}^T))$ and $\boldsymbol{P}^T \in L(N(\boldsymbol{R}^T))$, the claim follows.

ii) $\boldsymbol{T} = \boldsymbol{PS}$.

**Theorem 5** *For any lattice basis $\boldsymbol{R}$ of a sublattice of $L(N(\boldsymbol{A}))$*

$$X = X^{\boldsymbol{P},\boldsymbol{PS}}$$
$$Y = Y^{\boldsymbol{P},\boldsymbol{PS}}.$$

**Proof:** We show that $X^{\boldsymbol{I},\boldsymbol{Q}} \subseteq X^{\boldsymbol{P},\boldsymbol{PS}} \subseteq X$. If $\boldsymbol{x} \in X^{\boldsymbol{I},\boldsymbol{Q}}$, then there exists $\boldsymbol{\mu}$ such that $(\boldsymbol{x}, \boldsymbol{\mu}) \in U^{\boldsymbol{I},\boldsymbol{Q}}$, and then multiplying through by $\boldsymbol{P}$, we see that $(\boldsymbol{x}, \boldsymbol{\mu}) \in U^{\boldsymbol{P},\boldsymbol{PS}}$ as $\boldsymbol{PR} = \boldsymbol{0}$ and thus $\boldsymbol{x} \in X^{\boldsymbol{P},\boldsymbol{PS}}$. If $\boldsymbol{x} \in X^{\boldsymbol{P},\boldsymbol{PS}}$, then left multiplying by $\boldsymbol{M}$ gives $\boldsymbol{Ax} = \boldsymbol{Ax}^0 = \boldsymbol{b}$ as $\boldsymbol{MPS} = \boldsymbol{AS} = \boldsymbol{0}$, and thus $\boldsymbol{x} \in X^{\boldsymbol{A},\boldsymbol{0}}$. As $X^{\boldsymbol{I},\boldsymbol{Q}} = X^{\boldsymbol{A},\boldsymbol{0}}$, equality holds for all $\boldsymbol{P}$.

$Y^{\boldsymbol{I},\boldsymbol{Q}} = Y^{\boldsymbol{P},\boldsymbol{PS}}$ follows immediately from Lemma 1 as $\boldsymbol{P}^T$ is a basis for $L(N(\boldsymbol{R}^T))$, and $Y^{\boldsymbol{I},\boldsymbol{Q}} = Y^{\boldsymbol{A},\boldsymbol{0}}$ follows from Lemma 1 as $\boldsymbol{A}^T$ is a basis of $L(N(\boldsymbol{Q}))$. □

**Example 1** Consider a set $X = \{\boldsymbol{x} \in \mathbb{Z}^5 | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$ of the form

$$\begin{array}{rrrrrcl} x_1 & -5x_2 & -4x_3 & +11x_4 & +5x_5 & = & \boldsymbol{a}^1\boldsymbol{x}^0 \\ -13x_1 & -2x_2 & -12x_3 & -11x_4 & +x_5 & = & \boldsymbol{a}^2\boldsymbol{x}^0 \\ & & x & \in & \mathbb{Z}^5. & & \end{array}$$

Given

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 2 & 1 \end{pmatrix} \text{ and } \boldsymbol{M} = \begin{pmatrix} 1 & 0 & 5 \\ -12 & -1 & 1 \end{pmatrix},$$

it is easily checked that $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{P}$.

Now as $\boldsymbol{T} = \begin{pmatrix} -5 \\ 61 \\ 1 \end{pmatrix}$ is a basis of $L(N(\boldsymbol{M}))$, it follows from Proposition 3 that

$$\begin{array}{rrrrrcll} 1x_1 & & +1x_3 & +1x_4 & & = & \boldsymbol{p}^1\boldsymbol{x}^0 & -5\mu \\ 1x_1 & +1x_2 & -1x_3 & +1x_4 & & = & \boldsymbol{p}^2\boldsymbol{x}^0 & +61\mu \\ & -1x_2 & -1x_3 & +2x_4 & +1x_5 & = & \boldsymbol{p}^3\boldsymbol{x}^0 & +1\mu \\ & x & \in & \mathbb{Z}^5, & \mu & \in & \mathbb{Z}^1 & \end{array}$$

is an extended formulation for $X$.

An alternative is to calculate a basis $\boldsymbol{Q}$ of $L(N(\boldsymbol{A}))$, namely

$$\boldsymbol{Q} = \begin{pmatrix} 1 & -1 & 13 \\ 0 & 2 & 16 \\ 0 & 1 & -25 \\ -1 & 0 & 7 \\ 2 & 3 & -22 \end{pmatrix}.$$

Taking $\boldsymbol{R}$ to consist of the first two columns, and $\boldsymbol{S}$ to be the last column, we have that $\boldsymbol{P}^T$ from above is a basis of $L(N(\boldsymbol{R}^T))$ and $\boldsymbol{P}\boldsymbol{S} = \begin{pmatrix} -5 \\ 61 \\ 1 \end{pmatrix}$, so we arrive at the same extended formulation via the formulation of Theorem 5. □

# 3 Knapsack sets replaced by two equations

Here we analyze the case $m = 1$ and $s = 1$ in more detail. Specifically we are interested in the sets $X$ and $Y$ when

$$\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2.$$

We suppose that $a_j > 0$ for all $1 \leq j \leq n$, that $\gcd(a_1, \ldots, a_n) = 1$ and that $M_1, M_2 > 0$. Since $\gcd(a_1, \ldots, a_n) = 1$, it follows that $\gcd(M_1, M_2) = 1$, which

implies the existence of integers $\boldsymbol{q} \in \mathbb{Z}^2$ with $M_1 q_1 + M_2 q_2 = 1$. In addition we add the condition that $\mathrm{HNF}(\boldsymbol{P}) = (\boldsymbol{I}, \boldsymbol{0})$. Note that if $\mathrm{HNF}(\boldsymbol{P}) = (\boldsymbol{D}, \boldsymbol{0})$ with $\boldsymbol{D} \neq \boldsymbol{I}$, then it suffices to multiply the equation system by $\boldsymbol{D}^{-1}$ and work with $\boldsymbol{P'} = \boldsymbol{D}^{-1}\boldsymbol{P}$.

From Proposition 3, we know that $X^{\boldsymbol{P},\boldsymbol{T}} = X$, where $U^{\boldsymbol{P},\boldsymbol{T}}$ is of the form

$$\begin{aligned} \boldsymbol{p}^1\boldsymbol{x} &= \boldsymbol{p}^1\boldsymbol{x}^0 + M_2\mu \\ \boldsymbol{p}^2\boldsymbol{x} &= \boldsymbol{p}^2\boldsymbol{x}^0 - M_1\mu \\ & \quad \boldsymbol{x} \in \mathbb{Z}^n, \mu \in \mathbb{Z}^1. \end{aligned}$$

Note also that we can take $\boldsymbol{p}^i\boldsymbol{x}^0 = q_i b$ for $i = 1, 2$ as $M_1 q_1 + M_2 q_2 = 1$. This follows because $\boldsymbol{a}\boldsymbol{x}^0 - \boldsymbol{b} = M_1(\boldsymbol{p}^1\boldsymbol{x}^0 - q_1 b) + M_2(\boldsymbol{p}^2\boldsymbol{x}^0 - q_2 b) = 0$ and $\gcd(M_1, M_2) = 1$.

In the rest of this section we add a nonnegativity constraint on $\boldsymbol{x}$. Thus we consider

$$\begin{aligned} X_+ &= X \cap \mathbb{Z}^n_+ = \{\boldsymbol{x} \in \mathbb{Z}^n_+ \mid \boldsymbol{a}\boldsymbol{x} = b\} \text{ and} \\ Y_+ &= Y \cap \mathbb{R}^n_+ = \{\boldsymbol{x} \in \mathbb{R}^n_+ \mid \boldsymbol{a}\boldsymbol{x} = b\}. \end{aligned}$$

We derive two results. The first concerns the width of the polyhedron $V_+^{\boldsymbol{P},\boldsymbol{PS}}$ in the direction $\mu$. The second uses the width to derive a lower bound on the Frobenius number, simplifying and generalizing the result of Aardal and Lenstra [4] that is valid under the assumptions that $\boldsymbol{p}^1 \in \mathbb{Z}_{>0}$ and $M_2 = 1$.

### 3.1 The integer width

The integer width of a rational polytope $P$ in the integer direction $\boldsymbol{d}$, $w_I(P, \boldsymbol{d})$, is defined as

$$w_I(P, \boldsymbol{d}) = \left\lfloor \max\{\boldsymbol{d}^T\boldsymbol{x} \mid \boldsymbol{x} \in P\} \right\rfloor - \left\lceil \min\{\boldsymbol{d}^T\boldsymbol{x} \mid \boldsymbol{x} \in P\} \right\rceil + 1\,,$$

and is equal to the number of parallel lattice hyperplanes in direction $\boldsymbol{d}$ that are intersecting $P$.

Our goal is to calculate the integer width of

$$V_+^{\boldsymbol{P},\boldsymbol{PS}} = \{(\boldsymbol{x}, \mu) \in \mathbb{R}^n_+ \times \mathbb{R} \mid \boldsymbol{p}^1\boldsymbol{x} - M_2\mu = q_1 b,\ \boldsymbol{p}^2\boldsymbol{x} + M_1\mu = q_2 b\}$$

To do this, let

$$\bar{V}_+^{\boldsymbol{P},\boldsymbol{PS}} = \{(\boldsymbol{x}, \mu) \in \mathbb{R}^n_+ \times \mathbb{R} \mid \boldsymbol{p}^1\boldsymbol{x} - M_2\mu = q_1,\ \boldsymbol{p}^2\boldsymbol{x} + M_1\mu = q_2\} \tag{2}$$

be the scaled down version of this polyhedron with $b = 1$.

Below we derive the values $\bar{z} = \max\{\mu \mid (\boldsymbol{x}, \mu) \in \bar{V}_+^{\boldsymbol{P},\boldsymbol{PS}}\}$ and $\underline{z} = \min\{\mu \mid (\boldsymbol{x}, \mu) \in \bar{V}_+^{\boldsymbol{P},\boldsymbol{PS}}\}$. Once we have the values $\bar{z}$ and $\underline{z}$ it will be straightforward to compute the width $w_I(V_+^{\boldsymbol{P},\boldsymbol{PS}}, \boldsymbol{e}^{n+1})$ as $\lfloor b\bar{z} \rfloor - \lceil b\underline{z} \rceil + 1$.

**Lemma 6** *Consider formulation $\bar{V}_+^{\boldsymbol{P},\boldsymbol{PS}}$ (2), and let $k = \arg\max\{i|p_i^1/a_i\}$ and $j = \arg\min\{i|p_i^1/a_i\}$. Then,*

$$\bar{z} = \frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2} \text{ and } \underline{z} = \frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2}.$$

**Proof:** We consider the linear program

$$\begin{aligned}
& \bar{z} = \max \mu \\
\text{s.t.} \quad & \boldsymbol{p}^1\boldsymbol{x} - M_2\mu = q_1 && (3)\\
& \boldsymbol{p}^2\boldsymbol{x} + M_1\mu = q_2 && (4)\\
& \boldsymbol{x} \in \mathbb{R}^n_+, \qquad \mu \in \mathbb{R}^1.
\end{aligned}$$

Let $\boldsymbol{\gamma}$ be the dual variables corresponding to constraints (3)-(4). The corresponding dual problem is:

$$\begin{aligned}
\bar{z} = \min \quad & q_1\gamma_1 + q_2\gamma_2 && (5)\\
\text{s.t.} \quad & p_i^1\gamma_1 + p_i^2\gamma_2 \geq 0, \quad 1 \leq i \leq n, && (6)\\
& -M_2\gamma_1 + M_1\gamma_2 = 1, && (7)\\
& \boldsymbol{\gamma} \in \mathbb{R}^2
\end{aligned}$$

From constraint (7) we obtain

$$\gamma_1 = \frac{M_1\gamma_2 - 1}{M_2}\,. \tag{8}$$

Substituting for $\gamma_1$ in constraint (6) yields

$$p_i^1(\frac{M_1\gamma_2 - 1}{M_2}) + p_i^2\gamma_2 \geq 0, \quad 1 \leq i \leq n\,.$$

Rewriting gives $\gamma_2(p_i^2 + p_i^1(M_1/M_2)) - p_i^1/M_2 \geq 0$ for $1 \leq i \leq n$, which in turn yields $\gamma_2 \geq \frac{p_i^1}{a_i},\ 1 \leq i \leq n$. We now obtain

$$\gamma_2 = \frac{p_k^1}{a_k}\,. \tag{9}$$

Finally, we substitute for $\boldsymbol{\gamma}$ in the dual objective function (5) using expressions (8) and (9) which yields the optimal dual objective value

$$\bar{z} = q_1\left(\frac{M_1(p_k^1/a_k) - 1}{M_2}\right) + q_2\frac{p_k^1}{a_k} = \frac{p_k^1}{a_k}\left(\frac{q_1M_1 + q_2M_2}{M_2}\right) - \frac{q_1}{M_2} = \frac{p_k^1}{M_2a_k} - \frac{q_1}{M_2}\,.$$

The calculation of $\underline{z}$ is almost identical. □

Immediately we obtain the integer width.

**Theorem 7**

$$w_I(V_+^{\boldsymbol{P},\boldsymbol{PS}}, \boldsymbol{e}^{n+1}) = \left\lfloor \frac{bp_k^1}{M_2 a_k} - \frac{bq_1}{M_2} \right\rfloor - \left\lceil \frac{bp_j^1}{M_2 a_j} - \frac{bq_1}{M_2} \right\rceil + 1\,,$$

*where the indices $j$ and $k$ are defined as in Lemma 6.*

Notice that the choice of a valid $\boldsymbol{q}$ does not influence the width $w(V_+^{\boldsymbol{P},\boldsymbol{PS}}, \boldsymbol{e}^{n+1})$. Suppose $\boldsymbol{q'} \in \mathbb{Z}^2$ satisfies $M_1 q_1' + M_2 q_2' = 1$. The set of all valid $(q_1, q_2)^T$ can be written as

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} q_1' \\ q_2' \end{pmatrix} + \lambda \begin{pmatrix} -M_2 \\ M_1 \end{pmatrix},$$

where $\lambda \in \mathbb{Z}$. Thus a different choice of $q_1$ just means that the whole interval $[\underline{z}, \bar{z}]$ is shifted by an integer amount $\lambda$.

Observe also that given the choice of $\boldsymbol{R}$, any lattice basis $\boldsymbol{P}$ of $L(N(\boldsymbol{R}^T))$ yields the same width. The multipliers $M_1, M_2$ as well as the values $q_1, q_2$ do however change with different choices of $\boldsymbol{P}$.

**Example 2** Consider the following input vector $\boldsymbol{a}$

$$\boldsymbol{a} = (12223, 12224, 36674, 61119, 85569).$$

This is instance `cuww1` from Cornuéjols et al. [7]. With

$$\boldsymbol{P} = \begin{pmatrix} -1 & 0 & 2 & -1 & 1 \\ 2 & 1 & 1 & 6 & 6 \end{pmatrix}.$$

and $(M_1, M_2) = (12225, 12224)$, we have that $\boldsymbol{a} = \boldsymbol{MP}$, $\mathrm{HNF}(\boldsymbol{P}) = (\boldsymbol{I}, \boldsymbol{0})$.

Now we apply Theorem 7 with right-hand side $b = 89,643,481$, which is the Frobenius number of $\boldsymbol{a}$. We have $j = 1,\ k = 3,\ q_1 = 1,\ q_2 = -1$. We obtain

$$\begin{aligned}
&w_I(V_+^{\boldsymbol{P},\boldsymbol{PS}}, \boldsymbol{e}^{n+1}) = \\
&\left\lfloor \frac{bp_k^1}{M_2 a_k} - \frac{bq_1}{M_2} \right\rfloor - \left\lceil \frac{bp_j^1}{M_2 a_j} - \frac{bq_1}{M_2} \right\rceil + 1 = \left\lfloor \frac{bp_3^1}{M_2 a_3} - \frac{bq_1}{M_2} \right\rfloor - \left\lceil \frac{bp_1^1}{M_2 a_1} - \frac{bq_1}{M_2} \right\rceil + 1 = \\
&\left\lfloor \frac{89643481}{12224} \left( \frac{2}{36674} - 1 \right) \right\rfloor - \left\lceil \frac{89643481}{12224} \left( \frac{-1}{12223} - 1 \right) \right\rceil + 1 = \\
&\lfloor -7333.00003 \rfloor - \lceil -7333.9999 \rceil + 1 = -7334 + 7333 + 1 = 0.
\end{aligned}$$

It follows that $U_+^{\boldsymbol{P},\boldsymbol{PS}} = \emptyset$. Applying branch-and-bound, and branching first on the $\mu$ variable, this infeasibility would immediately be apparent. This is not the case using branch and bound starting from the original formulation $X_+^{\boldsymbol{A},\boldsymbol{0}} = \{x \in \mathbb{Z}_+^n \mid \boldsymbol{a}\boldsymbol{x} = b\}$. In particular Cplex fails to prove infeasibility within 500 million nodes. □

A natural question is whether the integer width differs if we use a different member of the family of extended formulations. Consider the sets

$$V_+^{\boldsymbol{I},\boldsymbol{Q}} = \{(\boldsymbol{x},\boldsymbol{\mu}) \in \mathbb{R}^n_+ \times \mathbb{R}^{s+r} \mid \boldsymbol{I}\boldsymbol{x} = \boldsymbol{I}\boldsymbol{x}^0 + \boldsymbol{S}\boldsymbol{\mu}^S + \boldsymbol{R}\boldsymbol{\mu}^R\}$$

and

$$V_+^{\boldsymbol{P},\boldsymbol{PS}} = \{(\boldsymbol{x},\boldsymbol{\mu}) \in \mathbb{R}^n_+ \times \mathbb{R}^{s} \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}^0 + \boldsymbol{P}\boldsymbol{S}\boldsymbol{\mu}^S\}$$

as described in Theorem 5. Using Lemma 1, we have that

$$\text{proj}_{x,\mu^S} V_+^{\boldsymbol{I},\boldsymbol{Q}} = V_+^{\boldsymbol{P},\boldsymbol{PS}}.$$

Thus we have

**Proposition 8** *$w_I(V_+^{\boldsymbol{P},\boldsymbol{PS}}, \boldsymbol{d}) = w_I(V_+^{\boldsymbol{I},\boldsymbol{Q}}, (\boldsymbol{d},\boldsymbol{0}))$, where $\boldsymbol{d}$ is any integer cost vector over the $(\boldsymbol{x},\boldsymbol{\mu}^S)$ variables.*

In particular, when $m = 1$, $s = 1$, and $\boldsymbol{d}$ is the unit vector corresponding to the last column of $\boldsymbol{S}$, then $w_I(V_+^{\boldsymbol{P},\boldsymbol{PS}}, \boldsymbol{e}^{n+1}) = w_I(V_+^{\boldsymbol{I},\boldsymbol{Q}}, (\boldsymbol{e}^{n+1},\boldsymbol{0}))$.

### 3.2 A lower bound on the Frobenius number

The Frobenius number of $\boldsymbol{a}$, $F(\boldsymbol{a})$, is the largest integer value of $b$ such that $\boldsymbol{a}\boldsymbol{x} = b$ does not have a nonnegative integer solution. Wlog we choose $\boldsymbol{q}$ such that $|q_1| \leq M_2/2$. So, if $|q_1'| > M_2/2$, we can determine new valid values of $q_1, q_2$ such that $|q_1'| \leq M_2/2$ by identifying an appropriate value of $\lambda$. In this section we still assume that $\text{HNF}(\boldsymbol{P}) = (\boldsymbol{I},\boldsymbol{0})$.

**Theorem 9** *Let $\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2$ with $\boldsymbol{a}, M_1, M_2, \boldsymbol{p}^1, \boldsymbol{p}^2$ satisfying the assumptions given in the beginning of Section 3. Moreover, let $\underline{z}$, $\bar{z}$ and the indices $j$ and $k$ be as defined in Lemma 6.*

*If $(-M_2/2) \leq q_1 \leq 0$ and*

*1a)* $\frac{p^1_j}{a_j} > q_1$

*2a)* $\frac{p^1_k}{a_k} < M_2 + q_1$

*3a)* $(\frac{1-\bar{z}}{\bar{z}-\underline{z}})\underline{z} \notin \mathbb{Z}$

*then*

$$F(\boldsymbol{a}) \geq \frac{a_j a_k (M_2 + q_1) - p^1_k a_j}{p^1_k a_j - p^1_j a_k} - \frac{M_2}{\frac{p^1_j}{a_j} - q_1}. \tag{10}$$

*or if $0 < q_1 \leq M_2/2$ and*

*1b)* $\frac{p_j^1}{a_j} > -M_2 + q_1$

*2b)* $\frac{p_k^1}{a_k} < q_1$

*3b)* $(\frac{1+\underline{z}}{\bar{z}-\underline{z}})\bar{z} \notin \mathbb{Z}$,

*then*

$$F(\boldsymbol{a}) \geq \frac{a_j a_k(M_2 - q_1) + p_j^1 a_k}{p_k^1 a_j - p_j^1 a_k} + \frac{M_2}{\frac{p_k^1}{a_k} - q_1}\,.$$

**Proof:**

We have already determined the width of $\bar{V}_+^{\boldsymbol{P},\boldsymbol{PS}}$ in the direction of $\mu$ corresponding to $b = 1$ in the proof of Lemma 6. Specifically we have shown that $\mu$ lies in the interval $[I_j,\ I_k]$, where

$$I_j := \underline{z} = \frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2} \text{ and } I_k := \bar{z} = \frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2},$$

whose width is

$$D := I_k - I_j = \frac{a_j p_k^1 - a_k p_j^1}{M_2 a_j a_k} > 0.$$

Any integer right-hand side value $b = t$ for which the corresponding interval $[tI_j,\ tI_k]$ does not contain an integer is a lower bound on the Frobenius number $F(\boldsymbol{a})$. Below we will show that

$$t \geq \frac{a_j a_k(M_2 + q_1) - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{M_2}{\frac{p_j^1}{a_j} - q_1}$$

is such a lower bound in the case that $-M_2/2 \leq q_1 < 0$. A sketch of the proof for the case $0 < q_1 \leq M_2/2$ is given in Appendix 1.

If $q_1 \leq 0$, Assumptions 1a and 2a imply that $0 < I_j < I_k < 1$. Moreover, since $q_1 \geq -M_2/2$ we obtain $I_k \leq p_k^1/(M_2 a_2) + 1/2$. Let $s := \frac{1-I_k}{D}$. Notice that $1 - I_k > 0$ since $I_k < 1$. The interval $[sI_j,\ sI_k]$ has length $1 - I_k$. Notice that $sI_j \notin \mathbb{Z}$ due to Assumption 3 of the theorem. Define $\ell := \lfloor sI_j \rfloor$ and $s' := \ell/I_j$. The number $s'$ satisfies $s - \frac{1}{I_j} < s' < s$, and yields the interval $[I_j', I_k'] := [s'I_j,\ s'I_k]$, with $I_j'$ integral. The length of $[I_j', I_k']$ is less than the length $1 - I_k$ of $[sI_j,\ sI_k]$. Therefore, $[I_j', I_k' + I_k]$ has length less than 1, and since $I_j'$ is integral it follows that $(I_j', I_k' + I_k]$ does not contain an integer.

Now, define $s^* := \lfloor s' \rfloor + 1$ and the interval $[I_j^*, I_k^*] := [s^* I_j,\ s^* I_k]$. We have $I_j' < I_j^* \leq I_j' + I_j$ and $I_k' < I_k^* \leq I_k' + I_k$. The result that $[I_j^*,\ I_k^*]$ does not contain an integer follows from the observation that $(I_j', I_k' + I_k]$ does not contain an integer.

We finally observe that

$$s^* = \lfloor s' \rfloor + 1 > \lfloor s - \frac{1}{I_j} \rfloor + 1 \geq s - \frac{1}{I_j} - 1 + 1 = s - \frac{1}{I_j}\,,$$

so we can conclude that $s - \frac{1}{I_j} = \frac{1-I_k}{D} - \frac{1}{I_j}$ yields a lower bound on the Frobenius number $F(\boldsymbol{a})$. Rewriting $\frac{1-I_k}{D} - \frac{1}{I_j}$ results in the expression

$$\frac{1-I_k}{D} - \frac{1}{I_j} = \frac{1 - \frac{p_k^1}{M_2 a_k} + \frac{q_1}{M_2}}{\frac{a_j p_k^1 - a_k p_j^1}{M_2 a_j a_k}} - \frac{1}{\frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2}} = \frac{a_j a_k (M_2 + q_1) - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{M_2}{\frac{p_j^1}{a_j} - q_1}\,.$$

□

We notice the similarity with the expression for the lower bound on the Frobenius number derived by Aardal and Lenstra [4] for the case that $M_2 = 1$ and $\boldsymbol{p}^1 \in \mathbb{Z}^n_{>0}$. If we set $M_2 = 1$ and $q = 0$ in Expression (10) we obtain

$$\frac{a_j a_k - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{a_j}{p_j^1}.$$

The only difference in the two expressions is in the numerator of the first term, where we have $p_k^1 a_j$ instead of $2p_j^1 a_k$ in [4]. This is a result of a different choice of the number $s$ in the proof. In [4] $s$ was chosen as $s = (1 - 2I_j)/D$ under a constraint on the relationship between $I_j$ and $I_k$.

# 4 Computation

## 4.1 Using reduced bases to find structure

Aardal and Lenstra [4] considered knapsack instances in which the input vector $\boldsymbol{a}$ can be decomposed as $\boldsymbol{a} = M_1 \boldsymbol{p}^1 + M_2 \boldsymbol{p}^2$, and used the reformulation earlier suggested in [3], cf. formulation $X^{I,Q}$:

$$\boldsymbol{x} = \boldsymbol{x}^0 + \boldsymbol{Q}\boldsymbol{\mu}\,,$$

where $\boldsymbol{x}^0$ and $\boldsymbol{Q}$ is as described in Section 2. They observed that if $\boldsymbol{a}$ is long, $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ are short compared to $\boldsymbol{a}$, and if the basis $\boldsymbol{Q}$ is reduced, then the first $n-2$ basis vectors of $\boldsymbol{Q}$ are short, and the last, $(n-1)$st basis vector is long. This can be explained as follows. First we observe that the orthogonal complement of the plane spanned by the vectors $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ is equal to $N(\boldsymbol{P}) \subset N(\boldsymbol{a})$. If $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ are short, the lattice $L(N(\boldsymbol{P}))$ contains short vectors yielding a relatively small lattice determinant. The rank of $L(N(\boldsymbol{P}))$, which is a sublattice of $L(N(\boldsymbol{a}))$, is just one less than the rank of $L(N(\boldsymbol{a}))$. Moreover, the determinant of the lattice $L(N(\boldsymbol{a}))$ is equal to the length of the vector $\boldsymbol{a}$. So, the large value of $d(L(N(\boldsymbol{a})))$ mainly has to be contributed by the basis vector that is in $L(N(\boldsymbol{a}))$ but not in $L(N(\boldsymbol{P}))$. Since basis

reduction orders the basis vectors in nondecreasing order of length, up to a multiplier, this basis vector is the last one in a reduced basis.

In general, we can derive a suitable decomposition of $\boldsymbol{A}$, in case no such decomposition is known a priori, by using the following algorithm.

i) Derive a reduced basis $\boldsymbol{Q}$ of $L(N(\boldsymbol{A}))$, see (1) in Section 1.1.

ii) Suppose $\boldsymbol{Q}$ consists of $s$ long vectors and $r = n-m-s$ short ones. How to define "long" and "short" is up to the user. (If all vectors of $\boldsymbol{Q}$ are of approximately the same length we set $s = n - m$). We define $\boldsymbol{R}$ to be the set of short vectors of $\boldsymbol{Q}$ and $\boldsymbol{S}$ to be the set of long ones.

iii) Find a reduced basis $\boldsymbol{P}^T$ of $L(N(\boldsymbol{R}^T))$.

iv) Solve the system of equations $\boldsymbol{MP} = \boldsymbol{A}$, $\boldsymbol{M} \in \mathbb{Z}^{m\times(m+s)}$ to find the matrix of multipliers $\boldsymbol{M}$.

**Example 3** We consider the same instance as in Example 2 with

$$\boldsymbol{a} = (12223, 12224, 36674, 61119, 85569).$$

Here we show how its hidden structure can be uncovered.

A reduced basis of $L(N(\boldsymbol{a}))$ is equal to

$$\boldsymbol{Q} = \begin{pmatrix} 0 & -3 & -1 & 2059 \\ 1 & 1 & -3 & 157 \\ -1 & -1 & -1 & -3336 \\ -1 & 1 & 0 & 2687 \\ 1 & 0 & 1 & -806 \end{pmatrix}.$$

Here we observe that the last column of the reduced basis $\boldsymbol{Q}$ is much longer than the other columns. Taking $r = 3$ and $s = 1$, $\boldsymbol{R}$ will consist of the first three columns of $\boldsymbol{Q}$, and $\boldsymbol{S}$ will consist of the last column of $\boldsymbol{Q}$. A reduced basis $\boldsymbol{P}^T$ for the lattice $L(N(\boldsymbol{R}^T))$ is

$$\boldsymbol{P}^T = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 2 & 1 \\ -1 & 6 \\ 1 & 6 \end{pmatrix}.$$

The vector $(M_1, M_2) = (12225, 12224)$ solves $\boldsymbol{MP} = \boldsymbol{a}$. We can now write $\boldsymbol{a} = 12225\boldsymbol{p}^1 + 12224\boldsymbol{p}^2$, with $\boldsymbol{p}^1$ being the first row of $\boldsymbol{P}$, and $\boldsymbol{p}^2$ being the second row of $\boldsymbol{P}$. Note that the matrix $\boldsymbol{P}$ obtained is not unique. In Example 2 a closely related but different basis of $L(N(\boldsymbol{R}^T))$ is considered. □

### 4.2 Feasibility testing and quality of the Frobenius bound

We tested the quality of the extended formulations for different choices of $s$ on some instances of integer equality knapsacks and the Cornuéjols-Dawande market split problem. For all instances of both problem types we compute a reduced basis $\boldsymbol{Q}$ and a vector $\boldsymbol{x}^0$ as described in Section 1.1, and derive matrices $\boldsymbol{P}$ and $\boldsymbol{M}$ as described in Section 4.1.

The integer knapsack instances were taken from Aardal and Lenstra [4]. Instances prob1–4 are such that the vector $\boldsymbol{a}$ decomposes with short $\boldsymbol{p}^1$, $\boldsymbol{p}^2$, whereas for the instances prob11–14 the $\boldsymbol{a}$-coefficients, randomly generated from $U[10000, 150000]$, are of the same size on average as in prob1–4. Instances prob11–14 have no apparent structure, and the columns of a reduced basis $\boldsymbol{Q}$ of $L(N(\boldsymbol{A}))$ are of approximately the same length. We use the Frobenius number of the vector $\boldsymbol{a}$ as right-hand side coefficient for all knapsack instances. Instances prob1–4 have 8 variables and prob10–14 have 10 variables. For details of the instances, see [4].

The market split instances [6] are multiple row equality knapsack problems in $\{0, 1\}$-variables with $m$ rows and $n = 10(m-1)$ variables. The elements of $\boldsymbol{a}^i$ for each row $i$ are generated randomly from $U[0, 99]$, and the right-hand side coefficients are calculated as $b_i = \lfloor (\sum_{j=1}^{n} a_j^i)/2 \rfloor$. We generated two sets of market split instances with 4 constraints and 30 variables, and 5 constraints and 40 variables respectively.

Table 1: The number of branch-and-bound nodes needed to solve the various reformulations for the knapsack instances.

| | | | $U_+^{\boldsymbol{P},\boldsymbol{PS}}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Instance | orig | AHL | $s=1$ | $s=2$ | $s=3$ | $s=4$ | $s=5$ | $s=6$ | $s=7$ | $s=9$ |
| prob1 | $>100$ mill. | 1 | 59 | 15 | 3 | 3 | 1 | 1 | 1 | – |
| prob2 | $>100$ mill. | 3 | 23 | 7 | 3 | 1 | 1 | 1 | 1 | – |
| prob3 | $>100$ mill. | 13 | 37 | 29 | 5 | 7 | 11 | 9 | 5 | – |
| prob4 | $>100$ mill. | 3 | 13 | 5 | 1 | 1 | 1 | 1 | 1 | – |
| prob11 | 100,943 | 61 | 2237 | 7683 | 317 | 89 | 51 | 69 | 49 | 61 |
| prob12 | 160,783 | 93 | 10,981 | 1105 | 967 | 523 | 179 | 105 | 117 | 71 |
| prob13 | 188,595 | 91 | 10,205 | 12,261 | 239 | 321 | 35 | 57 | 39 | 59 |
| prob14 | 140,301 | 87 | 2443 | 627 | 689 | 389 | 283 | 115 | 105 | 87 |

In Tables 1–3 we report on the number of nodes used by the integer programming solver Xpress Version 16.01.01 [16] to solve the various reformulations. Column "orig" refers to the original formulation in $\boldsymbol{x}$-variables. Column "AHL" refers to the Aardal-Hurkens-Lenstra lattice reformulation in which the $\boldsymbol{x}$-variables have been removed from the formulation, i.e., the formulation $\{\boldsymbol{\mu} \in \mathbb{Z}^{n-1} \mid \boldsymbol{Q}\boldsymbol{\mu} \geq -\boldsymbol{x}^0\}$ in the knapsack case and the formulation $\{\boldsymbol{\mu} \in \mathbb{Z}^{n-m} \mid -\boldsymbol{x}^0 \leq \boldsymbol{Q}\boldsymbol{\mu} \leq 1 - \boldsymbol{x}^0\}$ in the market split case.

For formulations $U_+^{\boldsymbol{P},\boldsymbol{PS}}$ we report on results for different values of $s$. Notice that the formulations AHL and $U_+^{\boldsymbol{P},\boldsymbol{PS}}$ for $s = n - m$ are mathematically equivalent, but the $U_+^{\boldsymbol{P},\boldsymbol{PS}}$-formulations contain the $\boldsymbol{x}$-variables with the identity matrix as coefficients. Since the solver reacts differently to the presence of the redundant $\boldsymbol{x}$-variables, this leads to slight deviations in the number of enumeration nodes needed.

Instances prob1–4, which decompose in short $\boldsymbol{p}^1, \boldsymbol{p}^2$, are very difficult to tackle with branch-and-bound applied to the original formulation. The Frobenius numbers for these instances are also large, see Table 4. None of the instances could be solved within 100 million nodes. As could be expected, the $U_+^{\boldsymbol{P},\boldsymbol{PS}}$-formulation with $s = 1$, which is a formulation with the $\boldsymbol{x}$-variables and one variable $\mu$, is easy to solve and comparable to the AHL-formulation. In contrast, instances prob11–14 are solvable using the original formulation, mainly due to the smaller value of the right-hand side coefficients. Here, one could expect that we would need to set $s = n - m$ to see a noticeable improvement compared to the original formulation, but in fact even taking $s = 1$ reduces the number of enumeration nodes by at least an order of magnitude, and with $s$ around 5 we obtain results comparable to those obtained with the AHL-formulation.

Table 2: The number of branch-and-bound nodes needed to solve the various reformulations: CD-instances $4 \times 30$.

| | | | $U_+^{\boldsymbol{P},\boldsymbol{PS}}$-ref | | | | | |
|---|---|---|---|---|---|---|---|---|
| Instance | orig | AHL | $s = 1$ | $s = 5$ | $s = 10$ | $s = 15$ | $s = 20$ | $s = 26$ |
| $4 \times 30\_1$ | 157,569 | 281 | 124,695 | 71641 | 8033 | 1397 | 1021 | 607 |
| $4 \times 30\_2$ | 169,455 | 167 | 154,505 | 51989 | 3794 | 1487 | 610 | 535 |
| $4 \times 30\_3$ | 209,741 | 325 | 178,697 | 181,373 | 32,367 | 1831 | 1025 | 845 |
| $4 \times 30\_4$ | 202,513 | 199 | 156,047 | 4685 | 3583 | 829 | 493 | 9527 |
| $4 \times 30\_5$ | 115,173 | 311 | 73,151 | 17,201 | 1197 | 391 | 353 | 3135 |

For the market split instances, which have no clear structure of the $\boldsymbol{Q}$-matrix, we notice similar results to those obtained for the knapsack instances prob11–14. The algorithm of Section 4.1 prescribes $s = n - m$ for these types of instances. The computational results suggest that smaller values of $s$ already yield significant computational improvement.

In Table 4 we report on the value of the Frobenius number as well as the value produced by the lower bound given in Theorem 9. For instances prob1–4 the lower bound is of the same order of magnitude as the Frobenius number, whereas for instances prob11–14 the bound is off by an order of magnitude. The bound might be improved by a different choice of the value $s$ in the proof of the theorem.

Table 3: The number of branch-and-bound nodes needed to solve the various reformulations: CD-instances $5 \times 40$.

| | | | $U_+^{\boldsymbol{P},\boldsymbol{PS}}$-ref | | | | |
|---|---|---|---|---|---|---|---|
| Instance | orig | AHL | $s = 5$ | $s = 10$ | $s = 20$ | $s = 30$ | $s = 35$ |
| $5 \times 40\_1$ | $> 10,000,000$ | 5873 | $> 10,000,000$ | 3,144,737 | 160,701 | 32,507 | 32,099 |
| $5 \times 40\_2$ | $> 10,000,000$ | 1643 | $> 10,000,000$ | 2,821,042 | 128,707 | 30,302 | 12,734 |
| $5 \times 40\_3$ | $> 10,000,000$ | 7349 | $> 10,000,000$ | 8,264,955 | 86,483 | 28,491 | 25,541 |
| $5 \times 40\_4$ | $> 10,000,000$ | 6870 | $> 10,000,000$ | 1,854,280 | 70,949 | 19,616 | 16,557 |
| $5 \times 40\_5$ | $> 10,000,000$ | 6651 | $> 10,000,000$ | 7,805,023 | 1,107,713 | 35,989 | 36,897 |

Table 4: The value of the lower bound of the Frobenius number.

| Instance | $F(\boldsymbol{a})$ | lower bound on $F(\boldsymbol{a})$ |
|---|---|---|
| prob1 | 33,367,335 | 26,061,675 |
| prob2 | 14,215,206 | 10,894,273 |
| prob3 | 58,424,799 | 31,510,625 |
| prob4 | 60,575,665 | 56,668,034 |
| prob11 | 577,134 | 98,774 |
| prob12 | 944,183 | 113,114 |
| prob13 | 765,260 | 67,752 |
| prob14 | 680,230 | 60,476 |

# References


[1] K. Aardal, R. E. Bixby, C. A. J. Hurkens, A. K. Lenstra and J. W. Smeltink. 2000. Market split and basis reduction: Towards a solution of the Cornuéjols-Dawande instances. *INFORMS Journal on Computing* **12** 192–202.

[2] K. Aardal, C. Hurkens, and A. K. Lenstra. 1998. Solving a linear diophantine equation with lower and upper bounds on the variables. In R. E. Bixby, E. A. Boyd, and R. Z. Ríos-Mercado, editors, *Integer Programming and Combinatorial Optimization, 6th International IPCO Conference*, volume 1412 of *Lecture Notes in Computer Science*, pages 229–242, Springer-Verlag, Berlin.

[3] K. Aardal, C. A. J. Hurkens and A. K. Lenstra. 2000. Solving a system of diophantine equations with lower and upper bounds on the variables. *Mathematics of Operations Research* **25** 427–442.

[4] K. Aardal and A. K. Lenstra. 2004. Hard equality constrained integer knapsacks. *Mathematics of Operations Research* **29(3)** 724–738. Erratum: *Mathematics of Operations Research* **31(4)**, 2006, page 846.

[5] J. W. S. Cassels. *An Introduction to the Geometry of Numbers.* Classics in Mathematics. Springer-Verlag, Berlin, 1997. Second Printing, Corrected, Reprint of the 1971 ed.

[6] G. Cornuéjols and M. Dawande. 1999. A class of hard small 0-1 programs. *INFORMS Journal on Computing* **11** 205–210.

[7] G. Cornuéjols, R. Urbaniak, R. Weismantel, and L. A. Wolsey. 1997. Decomposition of integer programs and of generating sets. R. E. Burkard, G. J. Woeginger, eds., *Algorithms – ESA '97.* Lecture Notes in Computer Science **1284**, Springer-Verlag, Berlin, Heidelberg, Germany, 92–103.

[8] R. Kannan. 1987. Algorithmic geometry of numbers. *Annual Review of Computer Science*, **2** 231–267.

[9] R. Kannan and A. Bachem. 1979. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. *SIAM Journal on Computing* **8** 499–507.

[10] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. 1982. Factoring polynomials with rational coefficients. *Mahematische Annalen* **261** 515–534.

[11] Lenstra, H. W., Jr. 2000. Flags and lattice basis reduction. C. Casacuberta, R. M. Miró-Roig, J. Verdera, S. Xambó-Descamps, eds., *Proceedings of the third European Congress of Mathematics Volume I*, Birkhäuser Verlag, Basel, 37–51.

[12] H. W. Lenstra, Jr. 2005. Lattices. To appear in *Surveys in algorithmic number theory*, Mathematical Sciences Research Institute Publications, Cambridge University Press, Cambridge, UK.

[13] L. Lovász. 1986. An Algorithmic Theory of Numbers, Graphs and Convexity. CBMS-NSF Regional Conference Series in applied mathematics **50** SIAM, Philadelphia, PA, USA.

[14] Q. Louveaux and L. A. Wolsey. 2002. Combining problem structure with basis reduction to solve a class of hard integer programs. *Mathematics of Operations Research*, **27(3)** 470–484.

[15] A. Schrijver. 1986. *Theory of Linear and Integer Programming.* John Wiley & Sons, Chichester, UK.

[16] Xpress-MP Optimization Software. Dash optimization. `http://www.dashoptimization.com/home/index.html`.

# Appendix 1

**Proof of Theorem 6 for the case $0 < q_1 \leq (M_2/2)$.**
If $0 < q_1 \leq (M_2/2)$, Assumptions 1b and 2b imply that $-1 < I_j < I_k < 0$, so the interval $[I_j,\ I_k]$ does not contain an integer. In addition, $I_j \geq p_j^1/(M_2 a_j) - 1/2$.

Let $s := \frac{1+I_j}{D}$. The length of the interval $[sI_j,\ sI_k]$ is equal to $1 + I_j$, and since $-1 < I_j < 0$ we have that $0 < 1 + I_j < 1$.

Notice that $sI_k \notin \mathbb{Z}$ due to Assumption 3b of the theorem. Define $\ell := \lceil sI_k \rceil$ and $s' := \ell/I_k$. The number $s'$ satisfies $s + \frac{1}{I_k} < s' < s$, and yields the interval $[I'_j, I'_k] := [s'I_j,\ s'I_k]$, with $I'_k$ integral. The length of $[I'_j, I'_k]$ is less than the length $1 + I_j$ of $[sI_j,\ sI_k]$. Therefore, $[I'_j + I_j, I'_k]$ has length less than 1, and since $I'_k$ is integral it follows that $[I'_j + I_j, I'_k)$ does not contain an integer.

Now, define $s^* := \lfloor s' \rfloor + 1$ and the interval $[I_j^*, I_k^*] := [s^* I_j,\ s^* I_k]$. We have $I'_j + I_j \leq I_j^* < I'_j$ and $I'_k + I_k \leq I_k^* < I'_k$. The result that $[I_j^*,\ I_k^*]$ does not contain an integer follows from the observation that $[I'_j + I_j, I'_k)$ does not contain an integer.

We finally observe that

$$s^* = \lfloor s' \rfloor + 1 > \lfloor s + \frac{1}{I_k} \rfloor + 1 \geq s + \frac{1}{I_k} - 1 + 1 = s + \frac{1}{I_k}\,,$$

so we can conclude that $s + \frac{1}{I_k} = \frac{1+I_j}{D} + \frac{1}{I_k}$ yields a lower bound on the Frobenius number $F(\boldsymbol{a})$. Rewriting $\frac{1+I_j}{D} + \frac{1}{I_k}$ results in the expression

$$\frac{1+I_k}{D} + \frac{1}{I_k} = \frac{1 + \frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2}}{\frac{a_j p_k^1 - a_k p_j^1}{M_2 a_j a_k}} + \frac{1}{\frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2}} = \frac{a_j a_k (M_2 - q_1) + p_j^1 a_k}{p_k^1 a_j - p_j^1 a_k} + \frac{M_2}{\frac{p_k^1}{a_k} - q_1}\,.$$